P1028

A deep learning approach for multiple sclerosis lesion segmentation

S. Valverde[1], M. Cabezas[1], E. Roura[1], S. González[1], D. Pareto[2], J.C. Vilanova[3], L. Ramió-Torrentà[4], À. Rovira[2], A. Oliver[1], X. Llado[1]

[1]ATC, University of Girona, Girona, [2]Magnetic Resonance Unit, Dept of Radiology, Vall d'Hebron University Hospital, Barcelona, [3]Girona Magnetic Resonance Center, [4]Neuroimmunology and Multiple Sclerosis Unit, Neurology Department, Dr. Josep Trueta University Hospital, Biomedical Research Institute (IDIBGI), REEM, Red Española de Esclerosis Múltiple, Girona, Spain

Background: MRI is extensively used in the diagnosis and monitoring of multiple sclerosis (MS), due to its sensitivity detecting focal white matter lesions (WML) in time and space. Although expert manual annotations of lesions is feasible in practice, this task is both time-consuming and prone to inter-observer variability, leading to the development of a wide number of automated lesion segmentation methods.

Aim: To present a novel pipeline for automated WM lesion segmentation of MS patient images based on a cascade of two 3D patch-wise convolutional neural networks (CNN).

Methods: The proposed pipeline relied on a cascade of two identical 7-layer CNN architecture. The first CNN was trained to be more sensitive revealing possible candidate lesion voxels while the second one was trained to reduce the number of misclassified voxels coming from the first network. To evaluate the performance, we first analysed the accuracy of our model on the 45 scans (T1-w, T2-w, T2-FLAIR) composing the public MS lesion segmentation challenge MICCAI2008 dataset. The optimised model using the available 20 training images was employed to compute the output segmentation masks of the 25 testing images, submitting them for on-line evaluation by the organizers. Furthermore, our model was also evaluated on two in-house MS clinical datasets (35 and 25 images, T1-w, T2-FLAIR) from the same medical centre, where WML masks were semi-automatically delineated by an expert radiologist of the same centre.

Results: Our method is today the best ranked approach on the MICCAI2008 board (1st / 60) when using all the available input modalities (T1-w, T2-w and T2-FLAIR), and still in the top-rank (3rd / 60) without using T2-w images. On clinical MS data, our model was superior than state-of-the-art methods such as LST and SLS, exhibiting a significant increase in the accuracy (53% of Dice overlap vs 33% and 32% in SLS and LST) and correlation ($r \geq 0.97$) with the expected lesion volume.

Conclusion: Our architecture is consistent with different input image modalities and image datasets, showing a significant increase in the sensitivity and low number of false positives. Our architecture tends to learn well from small sets of data, which is very interesting in practice, given the difficulty to obtain manual

label annotations and the amount of available unlabelled MRI data. The obtained results are encouraging, yielding our CNN architecture closer to human expert inter-rater variability.

Disclosure

S. Valverde: nothing to disclose.
M. Cabezas: nothing to disclose.
E. Roura: nothing to disclose.
S. González-Villà: nothing to disclose.
D. Pareto: nothing to disclose.
J.C. Vilanova: nothing to disclose.